EXPONENTIAL CONVERGENCE IN FEEDFORWARD ADAPTIVE SYSTEMS WITHOUT PERSISTENT EXCITATION

D. S. Bayard[1], J. Spanos, and Z, Rahman

Jet Propulsion Laboratory
California Institute of Technology
48000ak Grove Drive
Pasadena, CA 91109

Abstract

Conditions are investigated for exponential convergence of the tracking error in feedforward adaptive systems without persistent excitation. Particular attention is paid to the adaptive gradient algorithm in the overparametrized case. A main result is that for a bounded periodic regressor, the tracking error converges exponentially without regard to parameter convergence or to the degree of overparametrizati on, These results remove the persistent excitation (PE) conditions and parameter convergence conditions previously thought necessary to ensure exponential tracking error convergence in this class of systems.

1 Introduction

Persistent excitation (PE) conditions which ensure parameter convergence in adaptive algorithms have been studied by many researchers. Early results can be found in Astrom and Bohlin [1] where the PE condition is expressed in terms of positive definiteness of the autocorrelation function formed from the regressor. Subsequently, Bitmead and Anderson [4] proved that parameter convergence is *exponential* when PE conditions are satisfied in the adaptive gradient algorithm and the normalized adaptive gradient algorithms, Explicit upper and lower bounds on the exponential response can be found in [13], A general discussion of the PE condition is given in [3] and an effort to unify many definitions can be found in [15],

One important consequence of exponential parameter convergence, is that the tracking error (which is linear in the parameter error) also converges exponentially, This relationship gives the (false) impression that exponential tracking error convergence requires the same stringent PE conditions as parameter convergence, Interestingly, there are several

[1] Corresponding author: David S. Bayard, Mail Stop 198-326, bayard@bert2.jpl.nasa.gov

indications to the contrary. Using an approximate linear *analysis,* Glover [6] indicated as early as 1977 that exponential convergence of the tracking error is possible in the feed-forward adaptive gradient algorithm with tap delay line basis functions, and sinusoidal excitation, without any conditions on parameter convergence, More recently, Johansson [9] used a complete end-to-end Lyapunov analysis to demonstrate exponential tracking error convergence (to a bounded set) for an MRAC algorithm without persistent excitation or parameter convergence,

Motivated by Glover's approach, this paper investigates exponential convergence of the tracking error in adaptive feedforward systems, without regard to parameter convergence, A main result of the paper is that *for* any *bounded periodic regressor, the tracking error associated with the adaptive gradient algorithm converges exponentially without any PE condition.* This result is important, because in many applications good tracking performance is required while it is not desirable or even possible to satisfy PE conditions [12], The new analysis extends Glover's result by removing the need to use a tap delay line basis in the analysis (in fact, any method which generates a periodic regressor can be used). Furthermore, the new analysis is *exact* for finite-length regressors, in contrast to Glover's analysis which is only approximate unless the regressor is infinitely long,

A numerical example is given to demonstrate the usefulness of the exponential convergence results in the context of adaptive harmonic noise cancellation, The new results are critical in the case where the disturbance frequencies are not known a-priori, since the regressor must be overparametrized and the usual PE properties are lost.

In Section 2 the adaptive gradient algorithm is reviewed. Convergence properties are reviewed in Section 3 and new results pertaining to exponential tracking error convergence are presented, A numerical example is given in Section 4, which verifies the exponential nature of the error convergence and accuracy of the analytic bounds,

2 Adaptive Gradient Algorithm

Let the $y(t) \in R^1$ and $x(t) \in R^N$, be known signals and assume there exists a constant parameter vector $w^o \in R^N$ such that,

$$y(t) = w^{oT} x(t) \tag{1}$$

for all $t > 0$, Uniqueness of w^o is not required (i.e., the system can be overparametrized). An estimate $\hat{y}$ of y is constructed as,

$$\hat{y} = w(t)^T x(t) \tag{2}$$

where $w(t)$ is an estimate of w^o, tuned in real-time using the adaptive gradient algorithm [10],

$$\dot{w} = \mu x(t) e(t) \tag{3}$$

with adaptation gain $\mu > 0$, The tracking error is defined as,

$$e(t) = y(t) - \hat{y}(t) \tag{4}$$

and the parameter error is defined as,

$$\phi(t) = w^o - w(t) \tag{5}$$

Using (1)(2)(4)(5), the tracking and parameter errors can be related as follows,

$$e = \phi^T x(t) \tag{6}$$

Assuming that the true parameter w^o does not vary with time, (i.e., $\dot{w}^o = 0$), it follows from (3)(5) that,

$$\dot{\phi} = \dot{w}^o - \dot{w} = -\mu x e = -\mu x x^T \phi \tag{7}$$

This equation characterizes the propagation of the parameter error.

3 Exponential Convergence Properties

It is convenient at this point to review a well-known stability argument, Define the Lyapunov function candidate,

$$V = \frac{1}{2}\phi^T \phi \tag{8}$$

Taking the derivative of (8) and using (l)-(7) yields,

$$\dot{V} = -\mu e \phi^T x = -\mu e^2 \leq 0 \tag{9}$$

This proves that ϕ remains bounded. If x is bounded, then from (6) the error e remains bounded. Furthermore, if $\dot{x}$ is bounded, then V is bounded, V is uniformly continuous, and Barbalat's lemma ([1 O], pg. 85, and 276), can be applied to ensure that $\lim_{t \to \infty} e = 0$. This well known argument ensures that the error converges to zero as desired.

While the above argument ensures that e converges to zero, it does not indicate *how fast* it converges, Additional conditions can be imposed which ensure exponential convergence of c to zero. For example, if $x(t)$ is a periodic function with period ?'., then it is well known (cf., [10][1 l]), that the existence of some $\beta_1, \beta_2 > 0$ such that the following PE condition holds,

$$\beta_1 I \leq M \leq \beta_2 I \tag{10}$$

where M is the correlation matrix,

$$M = \frac{1}{T_o} \int_0^{T_o} x(t) x(t)^T dt \tag{11}$$

ensures that the error e converges exponentially: specifically, there exists constants $c_0 \geq 0$, $\alpha > 0$ such that,

$$|e| \leq c_0 e^{-\alpha t} \tag{12}$$

Precise expressions can be found for α in terms of β_1, β_2, μ and T_0 (i.e., set $\delta = T_0$, $\alpha_1 = \beta_1 \cdot T_o$ and $\alpha_2 = \beta_2 \cdot T_0$ in Lemma A.1 of Appendix A). In the case where μ is small, these expressions simplify to,

$$\alpha \simeq \mu\beta_1 \tag{13}$$

It is shown in this paper that while the PE condition (10) is required for exponential parameter convergence, it is overly restrictive for exponential tracking error convergence, Specifically, the main result proves that for x bounded and periodic, the convergence of e to zero is generically exponential *without any* condition *on* M. For example, in the important case where M is singular, it will be shown that an exponential convergence rate of the form (13) still holds, but with β_1 replaced by the smallest *nonzero* eigenvalue of M.

The main result of this paper is given next.

THEOREM 1 Assume there exists a $w^o \in R^N$ such that (1) holds for all $t \geq 0$, and that the adaptive gradient algorithm (2)-(7) is used to tune w, giving the following error system,

$$e = \phi^T x \tag{14}$$

$$\dot{\phi} = -\mu x x^T \phi \tag{15}$$

Let $x(t) \in R^N$ be a bounded periodic function of $t \geq 0$, with period T_0, i.e.,

$$||x(t)|| \leq \eta < \infty; \text{ for all } t \geq 0 \tag{16}$$

$$x(t + T_0) = x(t); \text{ for all } t \geq 0 \tag{17}$$

and let the eigenvector decomposition of its correlation matrix be defined as,

$$M = \frac{1}{T_o}\int_0^{T_o} x x^T d\tau = P \begin{bmatrix} \Lambda_{11} & 0 \\ 0 & 0 \end{bmatrix} P^T \tag{18}$$

$$\Lambda_{11} = diag\{\lambda_1, ..., \lambda_n\} > 0 \tag{19}$$

where $P^T = P^{-1}$, and, $\lambda_1 \geq ... \geq \lambda_n > 0$,

Then,

(i) The error system (14)(15) can be written equivalently as the reduced system,

$$e = r_1^T z_1 \tag{20}$$

$$\dot{r}_1 = -\mu z_1 z_1^T r_1 \tag{21}$$

$$\lambda_n T_o \cdot I \leq \int_0^{T_o} z_1 z_1^T d\tau \leq \lambda_1 T_o \cdot I \tag{22}$$

where,

$$z = P^T x; \qquad r = P^T \phi \tag{23}$$

$$z(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}; \quad r(t) = \begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix} \tag{24}$$

and $z_1, r_1 \in R^n$ and $z_2, r_2 \in R^{*-n}$. The smallest nonzero eigenvalue $\lambda_n > 0$ always exists unless $x \equiv 0$, $e \equiv 0$, i.e., the regressor x and error e vanish identically.

(ii) The tracking error e and reduced parameter error r_1 converge to zero exponentially as,

$$||r_1|| \leq c_0 e^{-\alpha t} \tag{25}$$

$$|e| \leq c_0 \eta e^{-\alpha t} \tag{26}$$

where,

$$\alpha = \frac{1}{2T_0} \ln \frac{1}{(1 - \alpha_3)} \tag{27}$$

$$c_0 = \left(\frac{1}{1 - \alpha_3}\right)^{\frac{1}{2}} \cdot ||r_1(0)|| \tag{28}$$

$$\alpha_3 = \frac{2\mu\lambda_n T_o}{(1 + \mu\lambda_1 T_o \sqrt{n})^2} \tag{29}$$

Letting μ be sufficiently small (i.e., such that $\mu << 1/(\lambda_1 T_o \sqrt{n})$), gives,

$$\alpha \simeq \mu\lambda_n \tag{30}$$

$$c_0 \simeq (1 + \mu\lambda_n T_o)||r_1(0)|| \tag{31}$$

PROOF:

Proof of (i): First consider the trivial case where $M = 0$. Then from (18), $x = 0$, and from (14) the error e vanishes identically.

Now consider the nontrivial case where $M \neq 0$. Then the eigenvalue decomposition (18) exists with some nonzero diagonal matrix $\Lambda_{11} > 0$. Using the transformed vectors z and r, the error equation (14) can be written as follows,

$$e = \phi^T x = \phi^T P P^T x = r^T z = r_1^T z_1 + r_2^T z_2 \tag{32}$$

Note that the error is simply the sum of the errors projected into two orthogonal subspaces. The correlation matrix of the transformed regressor z in (23) can be computed as,

$$\frac{1}{T_o} \int_0^{T_o} z z^T d\tau = P^T \left(\frac{1}{T_o} \int_0^{T_o} x x^T d\tau\right) P = \begin{bmatrix} \Lambda_{11} & 0 \\ 0 & 0 \end{bmatrix} \tag{33}$$

which implies from the partitioned structure of z in *(24)* that,

$$\frac{1}{T_o}\int_0^{T_o} z_1 z_1^T d\tau = \Lambda_{11} > 0 \qquad (34)$$

$$z_2(t) \equiv 0 \qquad \mathbf{(35)}$$

Substituting (35) into (32) gives the desired result (20). Note that the fact that z_2 is identically zero implies that the excitation in regressor $z = P^T x$ has been "squashed" completely into the smaller vector z_1 which by (34) is persistently exciting, Using (34), the correlation matrix of z_1 can be conveniently bounded from above and below as given in (22).

A similar reduction can be shown for the transformed parameter error r in (23). Specif- icall y, multiplying both sides of (15) on the left by P^T gives,

$$P^T \dot{\phi} = -\mu P^T x x^T \phi \qquad (36)$$

Substituting (23)(32)(35) into (36) gives,

$$\dot{r} = -\mu z z_1^T r_1 \qquad (37)$$

which can be partitioned using (24)(35) as,

$$\dot{r}_1 = -\mu z_1 z_1^T r_1 \qquad (38)$$

$$\dot{r}_2 = 0 \qquad (39)$$

Equation (38) is precisely *(21) as* desired.

Proof of (ii) Since the reduced regressor z_1 in (22) is persistently exciting, it follows that the reduced error system (20)(21) converges exponentially. In light of bounds in (22), Lemma A.1 of Appendix A can be applied with $\alpha_1 = \lambda_n T_o$, $\alpha_2 = \lambda_1 T_o$, and $\delta = T_o$ to give results (25)-(30) as desired, ∎

REMARK 1 Theorem 1 implies that the standard PE condition (10)(11) is overly re- strictive, In particular, if the regressor x is bounded and periodic, the convergence of the tracking error e to zero is generically exponential *without any condition on M.*

The most useful case is when *M* is singular for which the standard PE condition (10) fails to hold, yet a smallest nonzero eigenvalue $\lambda_n(M) > 0$ of $M = M^T$ *always* exists, thus determining the rate of exponential convergence by the new conditions (25)-(30). (The trivial case where $M = 0$ *is* uninteresting since the tracking error is identically zero).

REMARK 2 Intuitively, the persistent excitation conditions are eliminated in Theorem 1 by avoiding the need for convergence of the full parameter vector w in the proof, Rather, the "degree" to which the given regressor x is persistently exciting is indicated by the number n of nonzero eigenvalues of M. The parameter error vector ϕ is transformed and partitioned to become the vector r = $[r_1^T, r_2^T]^T$ such that $r_2 \in RN$-' is defined on a subspace which is not updated due to zero excitation, while in contrast $r_1 \in R^n$ is defined on a subspace which is excited persistently, Since the regressor z_1 associated with r_1 is persistently exciting, the reduced error vector r_1 converges exponentially, which from (20) ensures exponential convergence of e.

4 Numerical Example

An adaptive noise suppression example is given in this section demonstrating exponential tracking error convergence without persistent excitation, The set-up for the example in shown in Figure 1 using a Tap Delay Line with N taps, and delay T. Here, N = 50 parameters (i.e., taps) will be used to track m = **2** sinusoids. It is emphasized that this problem is heavily overparametrized since parameter convergence, according to the standard PE condition (10), requires at least N/2=25 sinusoids.

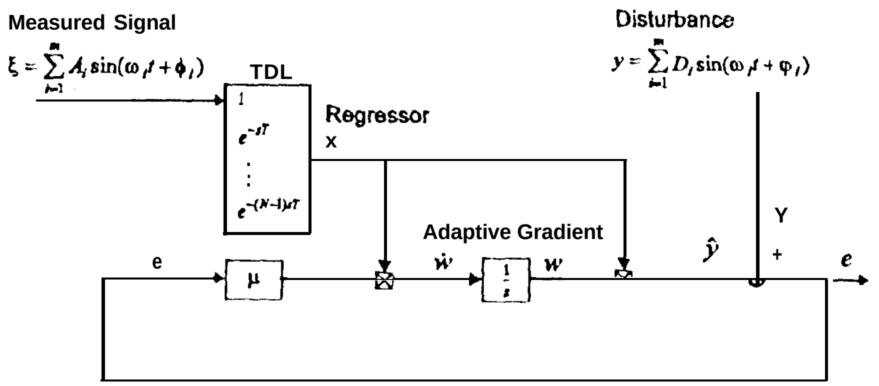

Figure 1: Configuration for adaptive noise suppression using a tap delay-line basis

Let y be a sum of two tones,

$$y = sin(\omega_1 t) + sin(\omega_2 t) \quad \textbf{(40)}$$

where $\omega_1 = 27r \cdot 25$, ω_2 = 27r , 50, and let the excitation signal ξ be given as,

$$\xi = A_1 sin(\omega_1 t - \pi/4) + A_2 sin(\omega_2 t - \pi/4) \quad (41)$$

where $A_1 = A_2 = \sqrt{2}$. Let the tap delay line basis be overparametrized with $N = 50$ taps and tap delay T = .004. The adaptive gradient algorithm (1)-(7) is used to tune the error to zero, using an adaptive gain of μ = ,1. A simulation of the response is shown as the solid line in Figure 2.

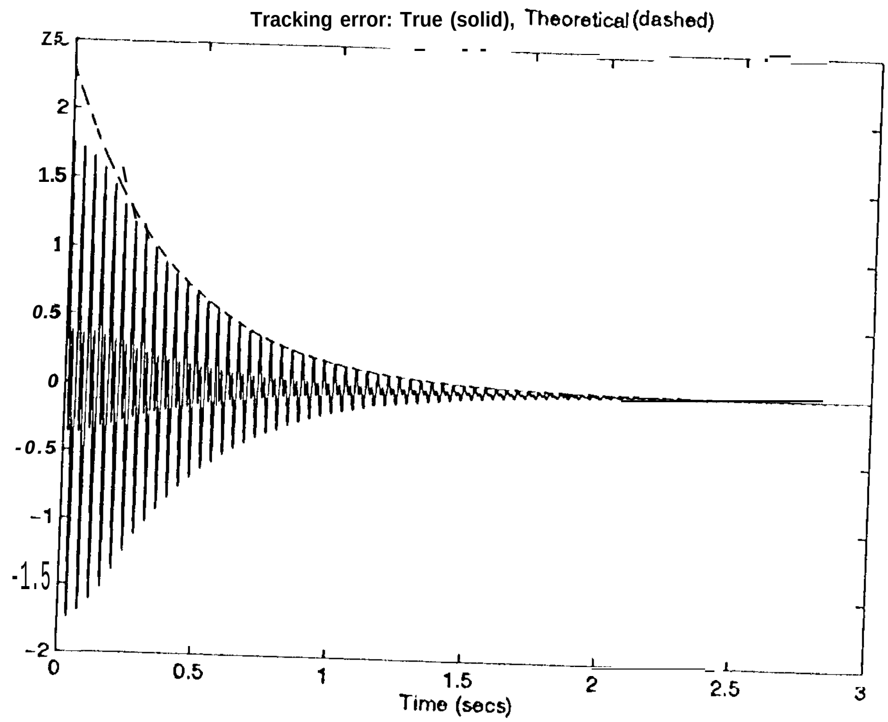

Figure 2: Exponential tracking error response from the adaptive gradient algorithm with heavily overparametrized tap delay line (N = 50 taps and m = 2 sinusoids).

The exponential convergence rate from Theorem 2 is calculated as $\alpha = 2.5000$, and the theoretical bound on e is shown as the dotted line in Figure 2. It is seen that the theoretical exponential overbound on the response is quite accurate,

In a 1977 paper [6], Glover studied the mapping from y to e in the LMS algorithm with a tap delay line basis, The LMS algorithm is the discrete-time equivalent of the adaptive gradient algorithm. Glover argues that for a sufficiently large number of taps, this mapping can be approximated as the linear time-invariant system, (specialized here to the case of two tones, and converted to continuous-time),

$$e = \frac{(s^2 + \omega_1^2)(s^2 + \omega_2^2)}{(s^2 + \omega_1^2)(s^2 + \omega_2^2) + \frac{\mu N}{2}\left(A_1^2(s^3 + \omega_2^2 s) + A_2^2(s^3 + \omega_1^2 s)\right)} \cdot y \qquad (42)$$

Note the nice interpretation of (42) as a stable double notch filter at precisely the frequencies of the disturbance. For the present example the closed-loop poles are located at, $(-2.4992 \pm 314.10j), (-2.5008 \pm 157.09j)$. Hence, Glover's approximate analysis predicts an exponential convergence with a rate which is determined by the real part – 2.4992 of the least damped pole, or equivalently α = 2.4992, This value agrees quite well with the value α = 2,5000 determined using the exact analysis of Theorem 2, which in turn agrees very well with the simulation results.

5 Conclusions

It is well known that PE conditions are required to ensure exponential convergence of the parameter error in adaptive systems, This coupled with a boundedness condition on the regressor ensures exponential convergence of the tracking error. However, in many applications one is only interested in the convergence of the tracking error. In such applications the PE conditions are unnecessarily stringent, In particular, the present paper shows that if the regressor is bounded and periodic, tracking error convergence in the gradient adaptation algorithm is generically exponential without regard to PE conditions.

In applications such as feedforward cancellation of harmonic disturbances (i.e., vibration damping, noise cancellation, etc.). the frequencies of the harmonic disturbances may be either known or unknown a-priori. When the disturbance frequencies are known (say there are m of them), it is often possible to construct a regressor of size $2m$ which satisfies the PE conditions directly, and hence ensures exponential convergence. However, in the more. common and useful case where the frequencies are not known a-priori, overparametrization is an essential tool to construct a regressor (of size larger that $2m$), which ensures a set of basis functions which is adequate over *the* entire *frequency range* of disturbances which may be encountered, Hence, the main results of this paper indicate that exponential convergence is retained in this important case, and motivates using overparametrization without loss of performance in such applications, For concreteness, a numerical example was given in which a regressor with 50 taps is used to track only 2 sinusoids, Standard PE conditions are clearly violated, yet the convergence is exponential with a rate which agrees quite nicely with the new theory.

In the present paper, the assumption that the regressor x is periodic is somewhat restrictive. Future efforts will be timed at relaxing this condition to include almost-periodic functions and possibly more generalized signals such as the Yuan-Wonham class [15].

Acknowledgements

This research was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

A APPENDIX A

LEMMA 1 *(Sastry and Bodson [11])* Consider the error equation,

$$e = \phi^T x \tag{1}$$

$$\dot{\phi} = -\mu x x^T \phi \tag{2}$$

where $\phi(t)$, $x(t) \in R^n$. Let x be a bounded piecewise continuous function of t such that,

$$||x(t)|| \leq \eta < \infty; \text{ for all } t \geq 0 \tag{3}$$

and let there exist constants $\alpha_1, \alpha_2, \delta > 0$ such that the following PE condition is satisfied,

$$\alpha_1 I \leq \int_t^{t+\delta} x(t)x(t)^T dt \leq \alpha_2 I, \text{ for all } t \geq 0 \tag{4}$$

Then the system **(1)(2)** is *globally exponentially stable*, i.e.,

$$||\phi|| \leq c_0 e^{-\alpha t} \tag{5}$$

$$|e| \leq c_0 \eta e^{-\alpha t} \tag{6}$$

where,

$$\alpha = \frac{1}{2\delta} \ln \frac{1}{(1-\alpha_3)} \tag{7}$$

$$c_0 = \left(\frac{1}{(1-\alpha_3)}\right)^{\frac{1}{2}} \cdot ||\phi(0)|| \tag{8}$$

$$\alpha_3 = \frac{2\mu\alpha_1}{(1+\mu\alpha_2\sqrt{n})^2} \tag{9}$$

Letting μ be sufficiently small (i.e., such that $\mu << 1/(\alpha_2\sqrt{n})$), gives,

$$\alpha \simeq \mu\alpha_1/\delta \tag{10}$$

$$c_0 \simeq (1 + \mu\alpha_1)||\phi(0)|| \tag{11}$$

PROOF: The proof follows directly from the development in Sastry and Bodson [11] pg. 73-?5 (see in particular Theorem 2.5.3) specialized to the gradient adaptation algorithm (1)(2).

References

[1] K.J. Astrom and T. Bohlin, "Numerical identification of linear dynamic systems from normal operating records," In *Theory of Self-Adaptive Control Systems.*, pp. 96-111, Plenum Press, New York, 1966.

[2] D.S. Bayard, J. Spanos and Z. Rahman, "Exponential convergence of the tracking error in adaptive systems without persistent excitation," IEEE Conf. on Decision and Control, New Orleans, LA, pp. 208-209, December 1995; Full version in JPL Internal Document, JPL D-12877, August 25, 1995.

[3] R.R. Bitmead, "Persistence of excitation conditions and the convergence of adaptive schemes," IEEE Trans. on Information Theory, vol. 30, pp. 183-191, March 1984.

[4] R.R. Bitmead and B.D.O. Anderson, "Lyapunov techniques for the exponential stability of linear difference equations with random coefficients," IEEE Trans. Automatic Control, vol. AC-25, pp. 782-787, August 1980,

[5] H. D'Angelo, *Linear Time- Varying Systems: Analysis and Synthesis.* Allyn and Bacon, Boston, 1970.

[6] J.R. Glover, Jr., "Adaptive noise canceling applied to sinusoidal interferences," IEEE Trans. Acoustics, Speech and Signal Processing, vol. 25, no. 6, December 1977.

[7] G .H. Golub and C.F. Van Loan, *Matrix Computations.* 2nd Edition, Johns Hopkins University Press, Baltimore Maryland, 1989,

[8] G .C. Goodwin and R.L. Payne, *Dynamic System Identification: Experiment design and Data Analysis,* Academic Press, New York, 1977.

[9] R. Johansson, "Global Lyapunov stability and exponential convergence of direct adaptive control," Int. J. Control, vol. 50, no. 3, pp. 859-869, 1989,

[10] K.S. Narendra and A.M. Annaswamy, *Stable Adaptive Systems.* Prentice-Hall, Englewood Cliffs, New Jersey, 1989,

[11] S. Sastry and M. Bodson, *Adaptive Control: Stability, Convergence, and Robustness.* Prentice Hall, 1989.

[12] J. Spanos and Z. Rahman, "Narrow-band control experiments in active vibration isolation," SPIE Proc., Vibration Monitoring and Control, paper 2264-01, San Diego, July 1994.

[13] A. Weiss and D. Mitra, "Digital adaptive filters: Conditions for convergence, rates of convergence, effects of noise and errors arising from implementation," IEEE Trans. on Information Theory, vol. IT-25, pp. 637-652, November 1979.

[14] B. Widrow and S.D. Stearns, *Adaptive Signal Processing.* Prentice-Hall, Englewood Cliffs, New Jersey, 1985.

[15] J. S-C Yuan and W.M. Wonham, "Probing signals for model reference identification," IEEE Trans. Automatic Control; vol. 22, pp. %30-538, August 1977,